

Chapter

Computational resources and strategies to construct single-molecule models of FISH

Beatriz Magalhães¹, Rita S. Santos¹, Nuno F. Azevedo¹, Anália Lourenço^{2,3,4,5}

¹ Laboratory for Process Engineering, Environment, Biotechnology and Energy (LEPABE), Department of Chemical Engineering, Faculty of Engineering of the University of Porto, Porto, Portugal

² Escuela Superior de Ingeniería Informática (ESEI), University of Vigo, Edificio Politécnico, Campus Universitario As Lagoas, Ourense, Spain

³ Centro de Investigaciones Biomédicas (CINBIO), University of Vigo, Campus Universitario Lagoas-Marcosende, 36310, Vigo, Spain

⁴ Sistemas Informáticos de Nueva Generación (SING) Research Group, Galicia Sur Health Research Institute (IIS Galicia Sur), SERGAS-UVIGO, Vigo, Spain

⁵ Centre of Biological Engineering (CEB), University of Minho, Campus de Gualtar, Braga, Portugal.

Abstract

Currently, the interactions occurring between oligonucleotides and the cellular envelope of bacteria are not fully resolved at the molecular level. Understanding these interactions is essential to gain insights on how to improve the internalization of the tagged oligonucleotides during fluorescence *in situ* hybridization (FISH). Agent-based modelling (ABM) is a promising *in silico* tool to dynamically simulate FISH and bring forward new knowledge on this process. Notably, it is important to simulate the whole bacterial cell, including the different layers of the cell envelope, given that the oligonucleotide must cross the envelope to reach its target in the cytosol. In addition, it is also important to characterise other molecules in the cell to best emulate the cell and represent molecular crowding. Here, we review the main information that should be compiled to construct an ABM on FISH and provide a practical example of an oligonucleotide targeting the 23S rRNA of *Escherichia coli*.

Running Head Construction of an agent-based model of FISH

Key words Fluorescence *in situ* hybridization, computational modelling, agent-based modelling, biological simulation, molecular diffusion.

1. Introduction

The reduced size of bacterial cells makes the direct observation by microscopy of its inner biological processes more complicated than for the larger animal cells. The construction of *in silico* models is a promising alternative to the use of experimental techniques. In particular, model simulations can be used to dynamically study *in vivo* behaviours and test different hypotheses, prior to engagement into wet lab work [3].

Fluorescence *in situ* hybridization (FISH), which is a largely used and well-known process for the detection of pathogens (in particular bacteria), can be a practical and relevant case study to model. Typically, FISH consists on a fluorochrome-tagged oligonucleotide which binds to a complementary RNA sequence of interest in the bacterial cytosol [2]. However, the cytosol is surrounded by the bacterial envelope, which is a major barrier for the penetration of oligonucleotides. There are still major knowledge gaps on the factors underlying such limited diffusion, most notably, regarding the diffusion fluctuations observed between bacteria of different groups [6]. Revealing this is key to uncover means to improve the internalization of oligonucleotides, and, consequently, boost the efficiency of FISH.

Among computational simulation tools, agent-based modelling (ABM) has successfully demonstrated the ability to simulate and observe various fine-grain aspects of biomolecular complexity [25, 31]. In this particular case, ABM may be a valuable tool to bring forward novel insights into the interaction of the oligonucleotides with the bacterial cell envelope and subsequent internalization.

Generally speaking, ABM represents the constituents of the system as autonomous decision-making entities, i.e. agents, and describes their main behavioural features based on theoretical and experimentally-derived rules [7]. Agent behaviour can range from the motion of a molecule to the binding of two molecules, and may adapt depending on the surrounding environment and the perceived situation [25]. In particular, by modelling the behaviour of individual agents as well as their social interplay, the behaviour of the whole system and emergent, more complex, phenomena can be observed [22]. The challenge is to represent the complexity of the biological system under analysis as faithfully as possible without complicating the model in excess and increasing computational costs [34]. In FISH, it is important to represent the entire cell, including the bacterial envelope, given that the tagged oligonucleotide must cross the envelope to reach its target in the cytosol. Fig. 1 summarizes the steps needed to undertake the ABM approach, from model construction to model validation.

In this chapter, we describe the implementation of such data workflow towards the construction of an ABM on FISH. Most notably, we describe the gathering of data and introduce available ABM frameworks and algorithms, pinpointing mandatory information towards both the construction and the simulation of such model. After a general description of the main methods, we introduce a case study of an oligonucleotide targeting the 23S rRNA of *Escherichia coli* (*E. coli*).

2. Methods

2.1 Literature Search

FISH involves multiple molecules interacting with each other and culminates in the detection of a fluorescence signal. This allows the detection and identification of bacterial cells, since they become fluorescent if the specific target sequence is present inside the bacterial cell. As a first modelling step, it is necessary to specify the most relevant agents involved in the system and compile data on their abundance, dimensions, shape, location within the cell (e.g. cytoplasmic membrane), and

behaviour (i.e. direction, speed, and interaction with other molecules), in order to be able to appropriately characterize them in the modelling environment.

Therefore, the proposed ABM considers the diffusion of molecules (in our case oligonucleotides) in the extracellular medium, in the cellular envelope of bacteria and within the cytoplasm. As such, the simulation environment, that is, the cell, including the bacterial cell envelope, must be characterized in terms of cell shape, dimensions, and volume.

2.1.1 Defining the simulation environment

The volume and dimensions of the cell are needed to adequately represent the simulation environment. Such parameters might be dependent on the strain, the phase in cell growth and even the growth conditions. So, the data curation workflow, combining literature and database searches, should account for:

1. The cell's dimensions, which depend on the cell's geometrical shape. If the cell is spherical in shape, the radius is enough to characterize the simulation environment; if the cell is cylindrical, both the length and diameter must be defined. Refer to Note 1 in case of missing data.
2. The thickness of each layer of the cell wall and the cytoplasmic membrane. Refer to Note 1 in case of missing data.
3. The dimensions of the cytoplasm, which can be obtained by subtracting the thicknesses of the membranes.
4. The definition of the extracellular volume, which should be sufficiently large to accommodate a statistically meaningful number of molecules.

2.1.2 Defining the Agents

The bacterial envelope and the cytosol contain thousands of molecular species but, due to the inherent high computational costs, biomolecular models typically only represent the most relevant molecules [12]. Therefore, the data curation workflow should proceed by:

1. Listing the oligonucleotide, fluorochrome, target sequence, and additional molecules, such as non-target nucleic acids, membrane phospholipids and proteins that might affect the oligonucleotide diffusion, which must be part of the modelled environment.
2. Performing independent queries for each molecule, looking for data on abundance, size, shape, volume, diffusion coefficient (Dc), and charge, if applicable. Refer to Notes 1, 2 and 3 in case of missing data. Furthermore:

- a. If the size is not found in the literature, calculate the hydrodynamic radius (Rh). Spherical approximation is typically assumed when representing a molecule in a computational simulation [10] and, in particular, the Rh is an effective approximation of its dimensions [17]. The Rh can be obtained as a function of the molecular weight (Mw) in Daltons. Use Equation 1 for RNA molecules, Equation 2 for proteins and Equations 3, 4 or 5 for DNA, depending whether its conformation is linear, coiled or supercoiled [18]. For some modified nucleic acid mimics (NAMs) and some proteins, an ellipsoid-like conformation is more realistic. In this case, use Equation 6 [28]. The volume of amino acids and sugars can be calculated using Equation 8 [36]. Typically, molecules are represented as being spherical or ellipsoid. As such, the radii of amino acids and sugars can be extrapolated from the calculated volume.

Equation 1 – RNA molecules

$$R_h(M_w) = 0.0515 \times M_w^{0.38}$$

Equation 2 – proteins

$$R_h(M_w) = 0.0515 \times M_w^{0.329}$$

Equation 3 – linear DNA

$$R_h(M_w) = 0.024 \times M_w^{0.57}$$

Equation 4 – coiled DNA

$$R_h(M_w) = 0.0125 \times M_w^{0.59}$$

Equation 5 – supercoiled DNA

$$R_h(M_w) = 0.0145 \times M_w^{0.57}$$

Equation 6 – ellipsoid molecules

$$R_h = aA(p)^{-1}$$

Where p is the ellipsoid aspect ratio ($p = b/a$), a is the measure of the polar semi axis in nm, and b is the measure of the equatorial semi axes in nm. A(p) can be calculated using Equation 7.

Equation 7

$$A(p) = \frac{1}{\sqrt{|1 - p^2|}} \begin{cases} \arctan \arctan \left(\sqrt{p^2 - 1} \right) & (\text{oblate: } p > 1) \\ \ln \ln \frac{1 + \sqrt{1 - p^2}}{p} & (\text{prolate: } p < 1) \end{cases}$$

Equation 8

$$V = \sum \text{all atom contributions} - 5.92N_B - 14.7R_A - 3.8R_{NA}$$

Where V is in Å³/molecule, the atom contributions are in cm³/mol, N_B is the number of bonds present in the molecule and can be calculated by Equation 9, R_A is the number of aromatic rings, and R_{NA} is the number of nonaromatic rings.

Equation 9

$$N_B = N - 1 + R_g$$

Where R_g is the total number of ring structures and can be calculated as follows (Equation 10):

Equation 10

$$R_g = R_A + R_{NA}$$

- b. If the Dc is not found in the literature, it can be calculated using the Stokes-Einstein equation (Equation 11) for spherical particles. Keep in mind that this Equation does not take into account molecular crowding. Nonetheless, it is the best approximation when experimental data is missing. For ellipsoid molecules, use Equation 12.

Equation 11 - Stokes-Einstein equation

$$D_c = \frac{K_b \times T}{6 \times \pi \times \eta \times R_h}$$

where K_b is the Boltzmann's constant with a value of $1.3806488 \times 10^{-23} \text{ m}^2 \cdot \text{kg} \cdot \text{s}^{-2} \cdot \text{K}^{-1}$, T is the temperature (estimated around 25 °C, equivalent to 298.12 K) and η is the viscosity of the medium in which the particle moves.

Equation 12 – Translational diffusion coefficient

$$tD_c = \frac{K_b T}{6 \times \pi \times \eta \times a} A(p)$$

3. Calculating the average time (τ) required for a molecule to move from one position in the system to another (Equation 13) [30]. This depends on the diffusion coefficient and on the size of the molecule.

Equation 13

$$\tau = \frac{\lambda^2}{6D_c}$$

where λ is the longest dimension of the molecule. We will use the R_h multiplied by two as the λ for spherical molecules and the equatorial axis as the λ for ellipsoid molecules.

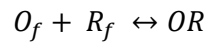
4. Adjusting agent populations according to the attainable total number of agents. Values of reference for molecule concentrations may be found in databases and papers. These values can be scaled up or down to agree with the volume defined for the simulation and, in particular, the number of agents should be adjusted in proportion to their abundance in the cell.

2.1.3 Defining the Interactions between agents

To represent the biological system as faithfully as possible, all interactions between the molecules of the system (i.e. the agents) should be described. Therefore, the data curation workflow should proceed by:

1. Describing all the behavioural rules for each agent, including hybridization (i.e. when the tagged oligonucleotide and ribosome collide with each other and become attached), dissociation (i.e. when a tagged oligonucleotide and a ribosome hybridized to each other detach from one another), rebound (i.e. when two molecules collide and their direction must be changed, given that the hybridization interaction does not apply), and movement (i.e. the default action carried out by the molecule to a vacant position if none of the other interactions is possible). Equation 14 represents the hybridization and dissociation interactions.

Equation 14 – Hybridization and Dissociation.



Where O_f represents a free oligonucleotide, R_f represents a free ribosome, OR represents an oligonucleotide attached (hybridized) to a ribosome.

2. Calculating the melting temperature (T_m) of the free oligonucleotide. The T_m can be defined as the temperature at which half of the oligonucleotides are hybridized and half are not, that is, the temperature at which the dissociation factor is 0.5. As such, the T_m can be used as an affinity parameter between the free oligonucleotide and the target sequence and a dissociation curve can be plotted, in order to get the dissociation factor at the temperature defined for the simulation. However, calculating the T_m is not straightforward and it depends on the type of oligonucleotide; the T_m can be calculated following the mathematical models described in [11]. For simplification, it can be assumed that whenever a collision between a free tagged oligonucleotide and a free ribosome occurs, they bind to each other. This assumption is only valid for a typical FISH temperature.

2.2 Simulation

Modelling cellular processes at the single-molecule level is currently feasible but extremely demanding in terms of computing power and software programming expertise. Hence, computational biologists often resort to coarser approaches to simulate large biological processes

[25]. In fact, most research facilities only have access to mid-performance computers and cannot realistically simulate highly-detailed models [26]. As a possible solution, the use of simulation frameworks can reduce some of the burden of implementing ABM. There are several frameworks which can be applied to model biological processes. These can either be general ABM platforms, such as Swarm [15], or Multi-Agent Simulation of Neighbourhoods (MASON) [21], or specific biologic simulators, such as ReaDDy [14] and Smoldyn [4]. The more general ABM frameworks allow the simulation of three-dimensional environments and of complex systems, involving thousands of heterogeneous agents, but demand skilled and cost-effective programming. Unlike the more general ABM toolkits, most of the specialized biomolecular simulation frameworks do not support three-dimensional modelling, and the ones which do require optimization or extension efforts. Therefore, choosing a modelling platform is not straightforward and should be carefully thought out, based on extensive surveys, experience and programming knowledge [26]. A review of the available simulation platforms is out of the scope of this work, as several reviews on the topic are already published [1, 19, 24].

To further reduce the demand in computing power, simpler simulation algorithms can be developed. In the single-thread approach the agents are introduced in a scheduler, which keeps track of all the agents in the model and executes them one by one. So, at each simulation step, each agent evaluates its neighbours and checks if any of its behavioural rules applies. In an attempt to create even more efficient and less-demanding algorithms, Pérez-Rodríguez *et. al* (2018), developed two alternative approaches [26]. The first alternative is the parallel approach, which introduces only one agent, named the controller, in the scheduler. The controller creates different threads that will handle the population of agents, that is, instead of iterating over one agent at a time, it iterates over multiple agents at the same time. In addition, it uses an optimized map-like structure that stores the agents by their position in one of three axes. The second approach, named the partitioning approach, also introduces only one agent in the scheduler. This agent creates several environment

partitions, each to be executed by a different thread and with its own map, which, again, decreases the number of iterations.

The simulation strategy should take into account the following decisions:

1. Choose the simulation framework, notably depending on the computational power at hand and familiarity with the required operating system and programming language.
2. Decide on the simulation approach to be implemented, namely single-thread, parallel, or spatially partitioned, or even, if desired, an in-house devised approach. Refer to Notes 4, 5, and 6 for additional information.
3. Set the initial number of agents and their initial positions (if not randomly assigned by the simulator).
4. Make sure that the simulator interprets model data correctly (e.g. the dimensions and diffusion of the agents, as well as main interaction rules).
5. Initiate the simulation.

2.3 Model Validation

Simulation outputs must be validated with experimental data. In the case of an ABM simulating FISH, the simulated diffusion times of the tagged oligonucleotide, both across the bacterial envelope and the cytoplasm, must be compared with experimental values, which can be obtained with techniques such as fluorescence recovery after photobleaching (FRAP) [32]. The optimized simulated concentration of uptake, that is, the minimum concentration, which gives out the maximum fluorescence signal, can also be compared with experimental values, by performing FISH and testing different concentrations. If the modelled diffusion times and the ideal concentration of uptake are statistically similar to the corresponding experimental values, the model is considered valid.

3. Practical Example on the construction of an ABM on FISH

Herein we present a practical example of the information that should be gathered to construct an ABM model of FISH. This model only considers one *E. coli* cell. Data was collected for an oligonucleotide typically used in FISH to target the 23S rRNA of *E. coli*, extracted from ProbeBase (accession no. pB-115 [9]).

The dimensions of *E. coli* K12, including the thickness of its cell envelope layers, are based on experimental values, notably a length of 2620 nm and a diameter of 680 nm [35]. Regarding the cell envelope, a lipid bilayer has a thickness of around 3 - 4 nm. In here, a thickness of 4 nm was considered for both the outer membrane and the cytoplasmic membrane, since a fine-grained model including detailed composition of each membrane is out of the scope of this tutorial. The periplasm has a thickness of 11-15 nm ([6] cited by [13]). In here, a thickness of 14 nm was considered. The dimensions of the cytoplasm can be calculated by subtracting the thickness of each layer of the cell envelope from the length and from the radius of the cell, which gives a cytoplasmic length of 2598 nm and a cytoplasmic radius of 658 nm. The model also accounts for the extracellular volume, where the tagged oligonucleotide is positioned at the start of simulation. The viscosity of the extracellular medium is assumed as equal to the viscosity of water (6.913×10^{-4} Pa/s at 37°C, [8]). Fig. 3 represents the simulation environment, modelled as a parallelepiped set to a length of 3144 nm and a width and height of 1032 nm. The dimensions of the parallelepiped were calculated by adding 20% to the length or width of the *E. coli* cell.

Table 1 contains the information about the ribosomes and its subunits, the oligonucleotide and the fluorophore, i.e. the agents of the model. All molecules were assumed to be spherical in shape, except for fluorescein and the tagged oligonucleotide, which were assumed to have an oblate shape. Fluorescein's was set with 0.7 nm and 0.2 nm as the equatorial and polar semi axes, respectively [27]. The equatorial axis of the tagged oligonucleotide was calculated as the sum of the equatorial axis of fluorescein and the diameter of the oligonucleotide, and the polar axis was assumed as equal to the diameter of the oligonucleotide. Moreover, its diffusion coefficient was calculated in the extracellular medium and in the cytoplasm, but not in each layer of the bacterial envelope, since

these are composed of heterogeneous molecules which form electrostatic interactions with each other and with the oligonucleotide.

To calculate the number of oligonucleotides in the extracellular medium at the beginning of the simulation, a concentration of 200 nM of oligonucleotides in hybridization solution was assumed. In addition, it was also assumed that 50 μ L of this solution is added to the solution containing the permeabilized bacteria (500 μ L). With this, the number of oligonucleotides in the considered extracellular volume can be extrapolated. Throughout the entire simulation, the same number of molecules should be present in the extracellular medium represented, which is achieved by simulating constant reposition of molecules, since in a FISH experiment there is an excess of oligonucleotides compared to the number of cells.

Regarding the interaction rules (Fig. 3), the following apply:

1. If a free tagged oligonucleotide collides with a free ribosome, both molecules stay attached to each other.
2. If two molecules, other than a free tagged oligonucleotide and a free ribosome, collide with each other, a rebound interaction occurs.
3. If a tagged oligonucleotide collides with the environment borders, a rebound interaction occurs.

4. Notes

Due to the incompleteness of data, especially of experimental diffusion values, and inconsistencies between sources, assumptions have to be made at this stage.

1. In the event that data is not available for the chosen microorganism, information can be obtained by phylogenetic comparison or following other criteria of similarity adequate for that particular information.
2. When data cannot be collected for a certain molecule, the search can be directed to other molecules with similar physicochemical characteristics and size.

3. Data should be scaled to the same units of measure. That is, do not use values in nm and μm in the same simulation.
4. If choosing a parallel or a spatially partitioned approach, select the number of threads, that is the number of agents the simulator should iterate at the same time.
5. If choosing a spatially partitioned approach, select the number of partitions, that is, the number of divisions of the environment.
6. When applying a partitioned or spatially partitioned approach, it is recommended to select at least as many threads as cores are available in the host machine, since the highest the number of cores, the more set of instructions can the machine process at the same time.

Acknowledgements

This work was financially supported by: project POCI-01-0145-FEDER-006939 (Laboratory for Process Engineering, Environment, Biotechnology and Energy – UID/EQU/00511/2013), funded by European Regional Development Fund (ERDF) through COMPETE2020 – Programa Operacional Competitividade e Internacionalização (POCI), by national funds (PIDDAC) through FCT – Fundação para a Ciência e a Tecnologia/MCTES; project “LEPABE-2-ECO-INNOVATION” – NORTE-01-0145-FEDER-000005, funded by Norte Portugal Regional Operational Programme (NORTE 2020), under PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF). Also received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 810685. Additionally, it was partially supported by the Consellería de Educación, Universidades e Formación Profesional (Xunta de Galicia) under the scope of the strategic funding of ED431C2018/55-GRC Competitive Reference Group, and the Portuguese Foundation for Science and Technology (FCT) under the scope of the strategic funding of UID/BIO/04469/2013 unit, COMPETE 2020 (POCI-01-0145-FEDER-006684), and PhD Grant SFRH/BD/143491/2019.

Fig. 1 – Basic data workflow of biomolecular agent-based modelling. Data retrieval, through literature curation and database (DB) queries, is used to gather the data required to describe the molecules and their interaction logic. The simulator is used to validate and execute the model. Firstly, the positions for each agent are set, either by the user or by random assignment, ensuring that for each agent the allocated position is free (i.e. not occupied by another agent). Once the initial positions are set, the simulator checks the environment, looking for possible interactions in the neighbourhood of each agent. Valid interactions are encoded in the behavioural rules. If no interaction rule is found, the simulator reallocates the agents and starts another iteration. The simulation terminates when the number of iterations or any of the monitored variables reaches a pre-established threshold. The modelling outputs should then be analysed and validated against experimental data through statistical comparison. If *in silico* results are in accordance with experimental results, the model is considered valid and may be used in new scenarios. If not, the model should be adjusted and re-evaluated.

Fig. 2 – Representation of the modelling environment and *E. coli* cell.

Fig. 3 – Interactions rules for the considered system. 1) A free tagged oligonucleotide collides with the environment border, leading to a rebound reaction (red arrow). 2) A free tagged oligonucleotide collides with a bound ribosome, leading to a rebound reaction of both molecules (red arrows). 3) A free tagged oligonucleotide collides with a free ribosome, leading to a binding reaction between both molecules.

Table 1 – Data to be introduced in the ABM representing FISH in *E. coli*.

Agent		Mw (KDa)	Rh (nm)	Dc (m ² /s)	V (nm ³)	τ (s)	Abundance (number/cell)	Length
50S Subunit	23S rRNA	990 [23]	9.78	1.33x10 ⁻¹⁵	3913.2 #	1.20x10 ¹⁶	18700 [23]	2904 nts [23]
	5S rRNA	41.0 [23]	2.92	4.45x10 ⁻¹⁵	103.8 #	3.18x10 ¹⁴	18700 [23]	120 nts [23]
	16S rRNA	520 [23]	7.65	1.70x10 ⁻¹⁵	1878.3 #	5.76x10 ¹⁵	18700 [23]	1541 nts [23]
Ribosome (70S)		-	21 [29]	6.18x10 ⁻¹⁶	2680 [33]	1.19x10 ¹⁷	14960 ###	-
Oligonucleotide		5.51 *	1.36	-	144.2 #	-	-	18 nts

Fluorescein	0.332 [16]	0.47	-	0.4 ^{##}	-	-	0.14 nm 0.4 nm
Tagged oligonucleotide	5.833 ^{**}	-	1,92x10 ⁻⁸ ; 2.67x10 ⁻¹⁵ ***	212.4 ^{##}	9.73x10 ¹⁴ *** ##	176 ^{\$}	-

Dc – diffusion coefficient; Mw – molecular weight; Rh – hydrodynamic radius; V – volume.

* Calculated with [20]. This calculus considers the Mw of the bases, the Mw of the modifications, the number of internal phosphates, the Mw of the phosphate groups, and the Mw of adjacent protons.

** Sum of oligonucleotide's Mw and fluorescein's Mw.

*** Calculated using the half of the sum of the diameter of the oligonucleotide and the axis of the fluorescein.

The volume was calculated using $\frac{4 \times \pi \times R h^3}{3}$

The volume was calculated using $\frac{4 \times \pi \times b^2}{3} \times a$

According to [5], approximately 80% of the 30S subunits are associated with the 50S, forming the complete 70S ribosome.

\$ Number of oligonucleotides in the extracellular medium.

References

1. Abar S, Theodoropoulos GK, Lemarinier P et al. (2017) Agent based modelling and simulation tools: a review of the state-of-art software. *Computer Science Review* 24:13-33
2. Amann R, Fuchs BM (2008) Single-cell identification in microbial communities by improved fluorescence in situ hybridization techniques. *Nature Reviews Microbiology* 6:339
3. An G, Mi Q, Dutta-Moscato J et al. (2009) Agent-based models in translational systems biology. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 1:159-171
4. Andrews SS, Addy NJ, Brent R et al. (2010) Detailed simulations of cell biology with Smoldyn 2.1. *PLoS computational biology* 6:e1000705
5. Bakshi S, Siryaporn A, Goulian M et al. (2012) Superresolution imaging of ribosomes and RNA polymerase in live *Escherichia coli* cells. *Molecular microbiology* 85:21-38
6. Beveridge T (1981) Ultrastructure, chemistry, and function of the bacterial wall. In: *International review of cytology*. Elsevier, p 229-317
7. Bonabeau E (2002) Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences* 99:7280-7287
8. Cooper J, Dooley R (2008) Release of the IAPWS formulation 2008 for the viscosity of ordinary water substance. *The International Association for the Properties of Water and Steam*
9. Daniel Greuter AL, Matthias Horn, and Thomas Rattei (2016) probeBase an online resource for rRNA-targeted oligonucleotide probes. In:
10. Feig M, Sugita Y (2013) Reaching new levels of realism in modeling biological macromolecules in cellular environments. *J Mol Graph Model* 45:144-156
11. Fontenete S, Guimarães N, Wengel J et al. (2016) Prediction of melting temperatures in fluorescence in situ hybridization (FISH) procedures using thermodynamic models. *Critical reviews in biotechnology* 36:566-577

12. Gameiro D, Pérez-Pérez M, Pérez-Rodríguez G et al. (2016) Computational resources and strategies to construct single-molecule metabolic models of microbial cells. *Briefings in bioinformatics* 17:863-876
13. Graham L, Beveridge T, Nanninga N (1991) Periplasmic space and the concept of the periplasm. *Trends in biochemical sciences* 16:328-329
14. Hoffmann M, Fröhner C, Noé F (2019) ReaDDy 2: Fast and flexible software framework for interacting-particle reaction dynamics. *PLoS computational biology* 15:e1006830
15. Iba H (2013) Agent-based modeling and simulation with Swarm. Chapman and Hall/CRC
16. Information NCFB Fluorescein. In:
17. Kalwarczyk T, Tabaka M, Holyst R (2012) Biologistics—diffusion coefficients for complete proteome of *Escherichia coli*. *Bioinformatics* 28:2971-2978
18. Kalwarczyk T, Tabaka M, Holyst R (2012) Biologistics - diffusion coefficients for complete proteome of *Escherichia coli*. *Bioinformatics* 28:2971-2978
19. Kravari K, Bassiliades N (2015) A survey of agent platforms. *Journal of Artificial Societies and Social Simulation* 18:11
20. Lgc OligoSpec Calculator. In:
21. Luke S, Cioffi-Revilla C, Panait L et al. (2005) Mason: A multiagent simulation environment. *Simulation* 81:517-527
22. Macal CM, North MJ (2010) Tutorial on agent-based modelling and simulation. *Journal of simulation* 4:151-162
23. Milo R, Jorgensen P, Moran U et al. (2010) BioNumbers - the database of key numbers in molecular and cell biology. *Nucleic acids research* 38:D750-D753
24. Nikolai C, Madey G (2009) Tools of the trade: A survey of various agent based modeling platforms. *Journal of Artificial Societies and Social Simulation* 12:2

25. Pérez-Rodríguez G, Dias S, Pérez-Pérez M et al. (2018) Agent-based model of diffusion of N-acyl homoserine lactones in a multicellular environment of *Pseudomonas aeruginosa* and *Candida albicans*. *Biofouling* 34:335-345
26. Pérez-Rodríguez G, Pérez-Pérez M, Fdez-Riverola F et al. (2016) High performance computing for three-dimensional agent-based molecular models. *Journal of Molecular Graphics and Modelling* 68:68-77
27. Pu Y, Wang W, Dorshow R et al. (2012) Picosecond polarization spectroscopy of fluorescein attached to different molecular volume polymer influenced by rotational motion. In: *Organic Photonic Materials and Devices XIV*. International Society for Optics and Photonics, p 825818
28. Roosen-Runge F, Hennig M, Zhang F et al. (2011) Protein self-diffusion in crowded solutions. *Proceedings of the National Academy of Sciences* 108:11815-11820
29. Schuwirth BS, Borovinskaya MA, Hau CW et al. (2005) Structures of the Bacterial Ribosome at 3.5 Å Resolution. *Science* 310:827-834
30. Siwiak M, Zielenkiewicz P (2010) A comprehensive, quantitative, and genome-wide model of translation. *PLoS computational biology* 6:e1000865
31. Troisi A, Wong V, Ratner MA (2005) An agent-based approach for modeling molecular self-organization. *Proceedings of the National Academy of Sciences* 102:255-260
32. Van Den Wildenberg SM, Bollen YJ, Peterman EJ (2011) How to quantify protein diffusion in the bacterial membrane. *Biopolymers* 95:312-321
33. Wittmann HG, Müssig J, Piefke J et al. (1982) Crystallization of *Escherichia coli* ribosomes. *FEBS Letters* 146:217-220
34. Yu J (2014) Coordination and Control Inside Simple Biomolecular Machines. In: Han K-I, Zhang X, Yang M-j (eds) *Protein Conformational Dynamics*. Springer International Publishing, Cham, p 353-384

35. Zaritsky A, Woldringh CL (1978) Chromosome replication rate and cell shape in *Escherichia coli*: lack of coupling. *Journal of bacteriology* 135:581-587
36. Zhao YH, Abraham MH, Zissimos AM (2003) Fast calculation of van der Waals volume as a sum of atomic and bond contributions and its application to drug compounds. *The Journal of organic chemistry* 68:7368-7373